Proceedings of the Twenty-Sixth Meeting of the Austronesian Formal Linguistics Association (AFLA)

Ileana Paul
(ed.)

AFLA XXVI
The Twenty-Sixth Meeting of the
Austronesian Formal Linguistics Association

University of Western Ontario
May 24-26, 2019

RELATIVE CLAUSE PROCESSING IN TONGAN: AN EFFECT OF SYNTACTIC ERGATIVITY ON THE OBJECT PREFERENCE

Hajime Ono, Tsuda University
Koichi Otaki, Kanazawa Gakuin Univ.
Manami Sato, Okinawa Int'l Univ.
'Ana Heti Veikune, USP, Tonga
Peseti Vea, USP, Tonga
Yuko Otsuka, Sophia University
Masatoshi Koizumi, Tohoku University

Table of Contents

PREFACE

The 26th Annual Meeting of the Austronesian Formal Linguistics Association (AFLA 26) was held on May 24-26, 2019 at the University of Western Ontario (Canada). The programme consisted of 24 presentations in addition to four plenary talks by Juliette Blevins, Vera Hohaus, Marian Klamer and Becky Tollan. This volume includes 13 papers from the conference.

As conference organizer, I received generous support from a variety of sources. Financial support came from the Social Sciences and Humanities Research Council of Canada (SSHRC), Research Western, the Joint Fund (Research Western, SOGS, SGPS), the Theoretical and Applied Linguistics Lab, the Canadian Linguistic Association, the Faculty of Arts and Humanities, the Graduate Program in Linguistics and three departments (French Studies, Modern Languages and Literatures, and Anthropology). The conference would not have been possible without the student volunteers (Sonia Masi, William Tran, Caylen Walker and Kang Xu), plus several others who helped out at the registration desk. Finally, I am grateful to the Department of French Studies for administrative support.

Many thanks to the abstract reviewers, to all those who attended, and to Mitcho Erlewine, who helped develop the current stylesheet.

Ileana Paul
University of Western Ontario

RELATIVE CLAUSE PROCESSING IN TONGAN: AN EFFECT OF SYNTACTIC ERGATIVITY ON THE OBJECT PREFERENCE*

Hajime Ono
Tsuda University
hajime@tsuda.ac.jp

Koichi Otaki
Kanazawa Gakuin Univ.
otaki@kanazawagu.ac.jp

Manami Sato
Okinawa Int'l Univ.
manamisato@gmail.com

'Ana Heti Veikune
USP, Tonga
ana.veikune@usp.ac.fj

Peseti Vea
USP, Tonga
peseta.vea@usp.ac.fj

Yuko Otsuka
Sophia University
yotsuka@sophia.ac.jp

Masatoshi Koizumi
Tohoku University
koizumi@tohoku.ac.jp

In this paper, we report our results obtained from a self-paced reading experiment in Tongan, a V-initial language with syntactic ergativity. Comparison of the reading time data from SRC and ORC suggests that native speakers of Tongan prefer the ORC structure, positing a gap in the transitive object position. This suggests that Tongan seems not to exhibit the subject advantage observed in other ergative languages like Avar, or other Austronesian languages like Chamorro. Instead, the data support that the absolutive case advantage based on the unmarkedness of the absolutive Case in ergative languages plays a major role in processing RCs. The lack of the subject advantage in Tongan may be linked to the syntactic ergativity of this language.

1. Introduction

One of the major goals in the study of human language processing is to propose an account that can handle structures in languages from a wide range of typological groups. In the literature, various hypotheses have been proposed, in order to account for native speakers' structural preferences, for example (Frazier and Rayner 1982, Crain and Fodor 1985, among others). However, the set of languages that have been investigated in those studies is quite limited, and typologically quite un-balanced

* We would like to thank the participants in the experiments as well as the staff members of the University of the South Pacific, Tonga Campus for their cooperation. We are also indebted to Dr. Raelyn Esau of the Tonga Ministry of Education and Training, without whose assistance we could not have conducted this research. We would like to express our sincere gratitude to the audience at AFLA 26 and the following people whose comments and criticism led to considerable improvement of this paper: Kentaro Nakatani, Noriaki Yusa and Masataka Yano. A previous version of this work has also been presented at Kansai Circle of Psycholinguists (Osaka, Japan). Of course, we are entirely responsible for remaining errors. Part of this work was supported by a Grant-in-Aid for Scientific Research (S) (#19H05589, PI: Masatoshi Koizumi) and (C) (#15K02529, PI: Hajime Ono) from the Japan Society for the Promotion of Science.

(Anand et al. 2011). Thus, it is necessary to examine to what extent some of those hypotheses proposed in the literature are truly "universal", and to investigate whether they are dependent on a certain grammatical profile of a given language (or a language family).

An example of the latter has already been reported in the literature, a well-known word order preference of placing Subject before Object (SO preference) depends on the syntactic structure of a given language (Koizumi et al. 2014, Koizumi and Kim 2016, Yano et al. 2017, Yasunaga et al. 2015). Those studies tested the word order preference in Kaqchikel, a Mayan language spoken in Guatemala and Truku Seediq, an Austronesian language spoken in Taiwan. Those languages have VOS as the basic word order, and they also allow SVO as a derived word order (García Matzar and Rodríguez Guaján 1997, Aldridge 2004). They observed through experiments that native speakers took more time to comprehend SVO sentences than VOS, and SVO sentences elicited a P600 effect, compared to VOS, indicating that the SVO structure involves a long filler-gap dependency which induces a greater processing cost. These findings suggest that the preference of placing subject before object is not universal.

In our current work we are interested in grammatical features such as verb-initial word order and ergative-absolutive Case system, both of which are known to be typologically rare (Dryer 2013, Comrie 2013). In (1), we have listed some psycholinguistic studies based on the languages with Verb-initial property or with Ergative case system.

(1) a. Verb-initial languages
Tzeltal (Norcliffe et al. 2015); Kaqchikel (see ref. above); Truku Seediq (Yano et al. submitted, Ono et al. submitted, Sato et al. submitted); Tagalog (Sauppe 2013, 2016, Tanaka 2016); Chamorro (Wagers et al. 2015, 2018)

b. Ergative languages
Basque (Carreiras et al. 2010); Avar (Polinsky et al. 2012); Ch'ol and Q'anob'al (Clemens et al. 2015); Niuean (Longenbaugh and Polinsky 2016, Tollan 2019); Georgian (Lau et al. 2018)

One of the goals in our current study is to add Tongan, a verb-initial language with ergative-absolutive Case system, to this set of languages.

In particular, we examined relative clause processing in Tongan, which we believe is suitable to shed some light on the universality of a certain processing preference. In many languages, it has been found that the processing cost for Subject Relative Clauses (or SRC) is lower than that for Object Relative Clauses (ORC). A typical paradigm is shown in (2). The processing advantage of SRC has been found in various measures such as reading time, eye-tracking, and so forth. We should note, however, that most of the previous studies on SRC advantage dealt with SVO or SOV languages, and typically with a nominative-accusative Case system, such as

English, German, French, Dutch, Korean, Japanese, Turkish, etc. (Kwon et al. 2013 for review).

(2) a. SRC: the doctor [who ___ criticized the nurse]
b. ORC: the doctor [who the nurse criticized ___]

Given that most of the psycholinguistic studies are not based on the verb-initial language nor ergative-absolutive Case system, it seems relevant to use Tongan in order to figure out whether an SRC advantage is universal, because Tongan has a very different grammatical profile compared to languages in which the SRC advantage has been observed. Therefore, in the current study, we investigate whether an SRC advantage is observed in Tongan.

In the next section, we review some grammatical properties in Tongan that are relevant for our experiment, then we go over some factors and hypotheses aiming at the SRC advantage found in many other languages. In section 4, we introduce our experiment, and section 5 concludes the paper.

2. Background: Tongan

In this section, we will review some grammatical properties of Tongan. As illustrated in (3) and (4), Tongan is a verb-initial language. Tongan has an ergative-absolutive case system, where the subject of intransitive verb and the object of transitive verb are marked with the same case called absolutive (Churchward 1953; Otsuka 2000). The subject of transitive verb is marked with a different case called ergative. The former is marked by a particle *'a* and the latter, by *'e*.[1]

(3) Na'e 'alu ['a e fefine] ki Tonga.
PST go ABS DEF woman to T
'The woman went to Tonga.'

(4) 'Oku 'ofa'i ['e Sione] ['a e fefine].
PRS love ERG J ABS DEF woman
'John loves the woman.'

Tongan relative clauses are post-nominal; relative clause follows the noun phrase (head) that it modifies as shown in (5).

[1] Strictly speaking, the article *e* (allomorph *he*) indicates specificity and not definiteness. The latter is expressed in Tongan phonologically as "definitive accent", stress on the final vowel of the final word of the relevant noun phrase, orthographically indicated as an acute accent, as in *fefiné* vs. *fefine*. In this paper, however, we gloss e/he as definite and dispense with orthographic representation of definitive accent in Tonga examples for the sake of simple exposition.

(5) Post-nominal RC (ORC, ABS-NP extracted)
'a e tōketā [na'e taa'i 'e he neesi ___]
ABS DEF doctor PST hit ERG DEF nurse
'the doctor who the nurse hit __ '

Relative clauses in Tongan exhibit syntactic ergativity (Otsuka 2000); (6) illustrates that when a 3rd person singular ergative NP is extracted for a relative clause, a resumptive pronoun *ne* must appear before the verb. (7), on the other hand, shows that such a requirement does not exist for the extraction of an absolutive NP.[2]

(6) SRC (ERG-NP extracted, RP required)
'a e tōketā [na'á ne taa'i ___ 'a e neesi]
ABS DEF doctor PST RP hit ABS DEF nurse
'the doctor who hit the nurse'

(7) SRC (ABS-NP extracted, intransitive verb)
'a e tōketā [na'e kata ___ mo e neesi]
ABS DEF doctor PST laugh with DEF nurse
'the doctor who laughed with the nurse'

This pronoun *ne* is also used for a 3rd person singular subject pronoun as shown in (8). This 3rd person singular subject pronoun may be optionally dropped.

(8) *ne* as a subject pronoun
Na'e taukave'i 'a e tōketā [na'á ne taa'i 'a e neesi].
PST claim ABS DEF doctor PST 3S hit ABS DEF nurse
'The doctor claimed that he hit the nurse.'

Thus, under a certain context, this pronoun *ne* creates a temporal ambiguity, between the resumptive pronoun and the regular subject pronoun, and this is an important grammatical feature of this language which we use in manipulating experimental sentences.

3. Processing of Relative Clauses

3.1. Major Accounts

Processing of relative clauses is one of the major topics in the literature of sentence processing (King and Just 1991, Traxler et al. 2002, Bader and Meng 1999, Kwon et al. 2013, among others), and as the SRC advantage is a well-known observation, a few major hypotheses have been proposed to account for the SRC advantage, arguing

[2] The past tense marker *na'e* has an allomorph *na'á*, which is used when it is followed by a clitic pronoun.

that the processing cost for SRC is lower than that for ORC. We will briefly review some of those accounts.

The Accessibility Hierarchy account (Keenan and Comrie 1977, 1979, Comrie and Keenan 1979) is proposed in order to explain a pattern of relative clause formation that a certain type of relative clause is more widely available than other types. Examining a number of languages, they propose that the relative clause formation of subject is more widely available than that of object, and that SRC is easier to comprehend than ORC, because subject is more prominent than object (Kwon et al. 2010, Polinsky et al. 2012).

Structural Distance hypothesis (Hawkins 2004; O'Grady 1997) suggests that SRC is easier than ORC because subject is structurally higher than object, and so the structural distance between the relative clause head noun and subject is shorter than the distance between the relative clause head noun and object. Assuming that such a dependency length affects the processing cost for a relative clause, this hypothesis accounts for the SRC advantage.

The last approach, the Dependency Locality Theory (Gibson 2000), argues that the processing cost is affected by the linear dependency length. In a post-nominal relative clause structure in English, for instance, the head noun of the relative clause appears before the relative clause. Given that subject comes before object, the subject position is linearly closer to the relative clause head. Then the processing cost for SRC is smaller than that for ORC.

Those approaches introduced above can account for the SRC advantage often observed in the literature. Although some accounts such the Dependency Locality Theory make wrong predictions with respect to the head-final relative clauses with SOV languages (like Japanese and Korean), each of those accounts relies on a very general notion or measure, which should be tested in a wide range of languages. In the next section we review some previous studies on relative clauses in languages with an ergative-absolutive Case system and / or in VSO languages, and see to what extent those general accounts can cover observations in those languages.

3.2. Previous Studies on Ergative and/or VSO languages

Although there are a number of studies investigating the comprehension process of relative clauses in the literature, one can find only a handful of studies on languages with ergative-absolutive Case system and/or verb-initial properties. Among those, we will briefly review three studies that seem quite relevant to the current study.

The first such study is by Carreiras et al. (2010) on Basque, an SOV language with prenominal relative clauses. Basque is known to have ergative-absolutive Case system, but it does not show syntactic ergativity; in other words, extraction of the ergative NP is equally allowed along with extraction of the absolutive NP for the relative clause formation. Through a self-paced reading experiment and ERP experiment, they observed that ORC was easier to process than SRC. They argue that morphological unmarkedness plays a role for the processing, which gives an advantage for the absolutive, which is null (then unmarked), over the ergative, which

is overtly marked by *-k* in Basque. They also argue that frequency cannot explain the ORC preference in Basque, showing that in a corpus, SRC is more frequent than ORC.

Next, Polinsky et al. (2012) examined relative clause processing of Avar, a Caucasian language. Avar is similar to Basque, in that it is an SOV language with a prenominal relative clause; it also exhibits a morphological ergative-absolutive Case system. They compared three types of relative clauses: intransitive (ABS) subject RC, transitive (ERG) subject RC, and transitive (ABS) object RC. Comparing the reading time of the head noun, they observed that intransitive (ABS) subject RC was read faster than the transitive RCs, and there was no reading time difference between the two transitive RC conditions. In order to explain the reading time pattern, they argue that both Case and the subject prominence (based on the Accessibility Hierarchy) are responsible for the processing cost. The Case factor here is, however, slightly different from what Carreiras et al. (2010) suggested, namely the morphological unmarkedness of the absolutive case marker. According to Polinsky et al. (2012), the ergative Case is a dependent Case in the ergative-absolutive Case system, in that it only appears when there is an absolutive case-marked NP in the same clause; on the other hand, the absolutive Case is an "independent" Case, because its appearance is not dependent on the existence of other NPs (Marantz 1991). They suggest that this "independent" property of the absolutive (or maybe as the default in the Case system) provides a processing advantage.[3]

The intransitive subject RC was processed fastest because the absolutive case-marked subject NP was relativized. Two transitive RC conditions receive some advantage from one of the two properties noted above, but crucially not both. The transitive (ERG) subject RC has an advantage with respect to the subject prominence, but the marked ergative Case was involved. The transitive (ABS) object RC has an advantage with respect to the unmarked status of the absolutive, but not on the subject prominence. Assuming that, somewhat metaphorically, the penalty for not satisfying the Case markedness is comparable to the penalty for not satisfying the subject prominence, they account for the observation that there was no difference between the two transitive RC conditions in terms of the processing cost .

Finally, there is a study on Chamorro, a verb-initial language of the Austronesian family like Tongan. Unlike Tongan, however, Chamorro is not an ergative-absolutive Case language. Also, Chamorro allows both prenominal and postnominal RCs. Wagers et al. (2018) showed that there was a strong subject-gap preference in Chamorro postnominal RCs, while there was a weak object-gap preference in prenominal RCs. This observation indicates that Chamorro exhibits the subject prominence probably in general, but at the same time, the linear distance

[3] Although Polinsky et al. (2012) did not discuss the morphological unmarkedness effect suggested by Carreiras et al. (2010), it might also be relevant for Avar. According to the description of the language in their paper, it seems that the absolutive does not have an overt case-marker, while the ergative has overt forms, such as *muradi-c:a* 'Muradi-ERG'. Therefore, the ergative case-marker is morphologically marked in Avar.

between the gap and the head noun of RC is also relevant. In the postnominal RCs, the subject-gapped RCs are "doubly" advantageous: they involve subject, and the subject-gap is linearly close to the head noun. On the other hand, in the prenominal RCs, the subject-gap seems to be preferred, but the position of the gap is relatively far from the head noun.

3.3. Research Questions

The above discussion tells us the following important points. Although languages with the ergative-absolutive Case system have been examined, those do not show syntactic ergativity, as far as we are aware of. Studies on Avar already show the unmarked case effect (Polinsky et al. 2012), then we expect to see a similar, or even stronger, effect in Tongan, which shows properties of syntactic ergativity.[4] Also, it should be noted that the processing cost of RCs is likely to reflect the combination of multiple factors. We have seen that factors interact, such as, the case markedness, the subject prominence, and the linear length between the gap position and the head noun, then we might see such an interaction in our current study. Therefore, we examine whether syntactic ergativity (i.e., syntactic markedness of ergative Case) plays a role in the processing of RCs, but at the same time, we pay attention to other factors such as the linear length and the subject prominence.

4. Experiment

This experiment was a self-paced reading experiment in which participants were asked to read sentences shown on the computer screen. The aim of this experiment was to examine the processing cost of relative clauses in Tongan; in particular, we are interested in comparing SRC and ORC. For that purpose, it is common in the literature on RC processing that SRC is directly compared against ORC. However, that would force us to compare RTs of NPs with different Case markers, as the template below illustrates.

(9) a. SRC. . . lawyer [$_{RC}$ PST ne chase <ERG-gap> ABS-pilot]
b. ORC. . . lawyer [$_{RC}$ PST chase ERG-pilot <ABS-gap>]

Because NPs with different Case markers may have different baseline reading times (possibly due to different markedness, etc.), we instead used a paradigm where each type of RCs is compared against the different control sentences. A sample set of the target sentences is shown here:

[4] Because the ergative case-marker in Tongan is no more complex than the absolutive case-marker (both being overt), it is hard to make a prediction in terms of the morphological markedness as in Basque (Carreiras et al. 2010).

(10) Sample set of target sentences

a. SRC Control (No gap) condition

Na‘e fakakaukau ‘a e tokotaha tā fakatātā na‘e tuli
PST think ABS DEF artist PST chase

‘e he loea ‘i he matātah ‘a e pailate koe‘uhi na‘e lolotonga
ERG DEF lawyer on DEF beach ABS DEF pilot because PST PROG

kikī hono kaume‘a.
scream his girlfriend

‘The artist thought that [the lawyer chased the pilot on the beach] because his girlfriend was screaming.’

b. SRC (Resumptive Pronoun) condition

Na‘e faitaa‘i ‘e he tokotaha tā fakatātā ‘a e loea na‘a ne
PST photograph ERG DEF artist ABS DEF lawyer PST RP

tuli ‘i he matātahi ‘a e pailate koe‘uhi na‘e lolotonga kikī
chase on DEF beach ABS DEF pilot because PST PROG scream

hono kaume‘a.
his girlfriend

‘The artist photographed the lawyer [who _ chased the pilot on the beach] because his girlfriend was screaming.’

c. ORC Control (Plural) condition

Na‘e tautea ‘e he kau fakamaau ‘a e loea na‘a nau
PST punish ERG DEF PL judge ABS DEF lawyer PST they

faka‘ita‘i ‘i he fakamaau‘angá neongo na‘á nau kaungāme‘a fuoloa.
anger in DEF court although PST they friends long time

‘The judges punished the lawyer [who they angered _ in the court] although they were friends for a long time.’

d. ORC (Singular) condition

Na‘e tautea ‘e he fakamaau ‘a e loea na‘á ne faka‘ita‘i
PST punish ERG DEF judge ABS DEF lawyer PST 3S anger

‘i he fakamaau‘anga neongo na‘á nau kaungāme‘a fuoloa.
in DEF court although PST they friends long time

‘The judge(s) punished the lawyer [who {s/he} angered _ in the court] although they were friends for a long time.’

In the SRC pairs (10ab), the absolutive case-marked NP (‘*a e pailate* “ABS-pilot”) in the embedded clause is the critical region. In (10a), the embedded clause is a complement clause and there is no sign of dislocated element. Given the

transitive verb (*tuli* "chase") in the embedded clause, the native speakers of Tongan will expect to see an absolutive case-marked NP. In (10b), the object in the matrix clause (*'a e loea* "the lawyer") is followed and modified by a relative clause. Although Tongan does not have a dedicated RC marker, the clause should be a relative clause because the matrix verb (*faitaa'i* "photograph") does not take a complement clause. If native speakers of Tongan prefer ORC, they should posit an ORC structure. This leads them to hypothesize the clitic *ne*, appearing before the verb, to be a subject pronoun, possibly referring to the matrix subject. Furthermore, crucially they should expect a gap in the object position. Under this expectation, they will be surprised to see the overt absolutive case-marked NP in the relative clause, because this NP shows the readers that their hypothesis was wrong, and they were actually seeing an SRC structure. This forces them to revise and rebuild the structure, which leads to a reading time slowdown. In the SRC sentence pairs, an adjunct clause involving *because*, *after*, *before*, etc. was added at the end in order to avoid any sentence-final wrap-up effect and to measure a spill-over effect.

In the ORC pairs (10cd), the number feature on the clitic pronoun in the relative clause was manipulated. In (10c), the plural clitic pronoun *nau* was used, and this pronoun refers to the plural matrix subject (*'e he kau fakamaaú* "ERG-the judges"), and cannot refer to the head noun, which is singular. Therefore, the relative clause should be an ORC structure in which the native speakers can expect an object gap. In (10d), the singular clitic pronoun *ne* was used. If native speakers of Tongan prefer SRC, they should take the clitic *ne* as a resumptive pronoun for the subject extraction, and expect to see an absolutive case-marked NP in the RC structure. The target sentence (10d) actually does not have such an NP, then they will be surprised to see an adjunct clause marker such as *neongo* "although", because it strongly suggests the end of RC. This adjunct clause marker is the critical region in the ORC pairs.

In sum, we examine in which pair, SRC or ORC, the participants show a slowdown compared to the baseline condition. If the ORC structure is preferred, the native speakers of Tongan will slow down by the presence of ABS-NP in (10b). On the other hand, if the SRC structure is preferred, they will slow down by the absence of ABS-NP in (10d), compared to the baseline.

4.1. Method

Participants in this experiment included 61 native speakers of Tongan (40 females, mean age = 24.2, SD = 8.0, range = 18-54). All of them were students or staff members, recruited at the University of the South Pacific, Tonga campus, and they were naïve to the purpose of the experiment. Written informed consent was obtained, as well as some linguistic background questionnaire given prior to the experiment. They were provided some food package for their participation of the experiment.

Sixteen sets of SRC pairs and sixteen sets of ORC pairs of target sentences were constructed as in (10). In addition to the target sentences, 40 filler sentences based on five different sentence types were also constructed. They were similar to

target sentences in terms of the length and complexity. All sentences used in the experiment was checked for their acceptability by two native speakers of Tongan.

Approval for the study was obtained from the Prime Minister's Office of Kingdom of Tonga, the Ethics Committee of the Graduate School of Arts and Letters, Tohoku University, Japan and the Research Ethics Committee of the University of the South Pacific.

4.2. Procedure

The target sentences were distributed into two lists in a Latin-squared design, so that each participant only sees one version in each pair (sixteen SRC-related sentences and sixteen ORC-related sentences). 40 filler sentences were added to the list so that there were 72 trials in total. An experimenter who is a native speaker of Tongan explained the task and procedures, and answered questions if any. Five practice items were presented before the main experiment in order for the participants to get familiarized with the task. Each participant sat in front of the laptop computer in a quiet room and silently read sentences on display. The experiment was run in a moving-window, non-cumulative style with the software Linger (developed by D. Rohde), measuring reading times for each phrase on display while the participants were instructed to press the space bar to read the following phrase at their normal reading rate (Just et al. 1982). For instance, (10d) was divided into multiple regions at the positions marked with '/' in the following, roughly corresponding to a phrase.

(10) d. Na'e/ tautea/ 'e he fakamaau/ 'a e loea/ na'á ne/ faka'ita'i/
PST punish ERG DEF judge ABS DEF lawyer PST 3S anger
'i he fakamaau'anga/ neongo/ na'á nau kaungāme'a/ fuoloa.
in DEF court although PST they friends long time

Each experimental sentence was followed by a comprehension question asking about the content of the sentence they had just read. The participants pressed either the F key (yes) or the J key (no) to answer the comprehension question. A feedback was given when they wrongly answered the question. The entire experiment lasted approximately 30-45 minutes in total including the instruction.

4.3. Data Analysis

Data from a few participants and target items are removed. First, we removed data from one participant and one target item whose mean accuracy rate for comprehension question was lower than 2 standard deviations away from the total mean accuracy rate. We analyzed the mean accuracy rates for each condition using ANOVA. Second, we removed data from two participants whose mean reading times were longer than 2 standard deviations away from the total mean reading times. Furthermore, we removed reading time data from the trials where their comprehension question was not correctly answered, as well as reading time data

longer than 5,000ms from any region. Based on the remaining reading time data, we performed the linear mixed effect regression analyses with the *lme4* package in R version 3.4.0 (R Development Core Team 2014, Baayen 2008, Baayen et al. 2008, Bates et al. 2014).

Because the SRC and ORC pairs are independent, we built separate models, using the sentence type as fixed factor, and random intercepts and random slopes were estimated for participants and items. We estimated a model first, and removed data whose residual was greater than 2.5 standard deviations (Baayen and Milin 2010), then we re-built the final model.

4.4. Results

Comparing the mean accuracy rates for each condition, we found no difference between the two SRC conditions (Control (10a), 73.0%, Resumptive Pronoun (10b), 72.5%), whereas the mean accuracy rate for the ORC Singular condition ((10d), 74.1%) was significantly lower than that for the ORC Control condition ((10c), 81.9%) ($F(1,57) = 5.61$, $p < 0.03$). Some might feel that the overall mean accuracy rates were slightly lower than those typically reported in the literature. We speculate that this is because some participants were very nervous and/or too cautious, for they were not so familiar with the situation of the experiment like ours.

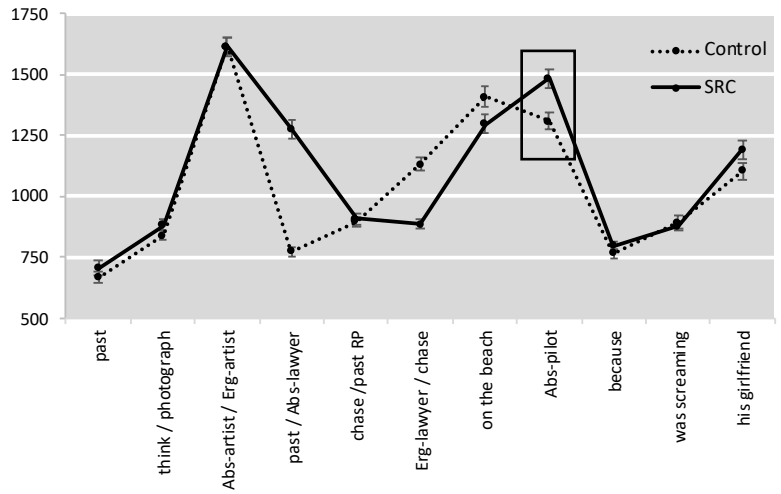

Figure 1. Region-by-region reading times for the SRC pairs.
The box indicates the critical region abs-NP. Error bars represent SEs.

Turning to the reading time data, Figure 1 shows region-by-region mean reading times for the SRC pairs, and Figure 2 for the ORC pairs. In the SRC pairs, there was a slowdown at the critical region (Region 8, ABS-NP) in SRC Resumptive Pronoun condition ($\beta = 183.17$, SE = 38.19, $t = 4.80$, $p < 0.001$). In the ORC pairs, on the other hand, there was no slowdown at the critical region (Region 8, an adjunct

clause marker) in the ORC Singular condition, compared to its Control condition. One complication of the data for the ORC pair is that there was a slowdown in Region 7 (PP) ($\beta = -126.58$, SE = 37.28, $t = -3.40$, $p < 0.001$). We provide some brief speculations in the next section.

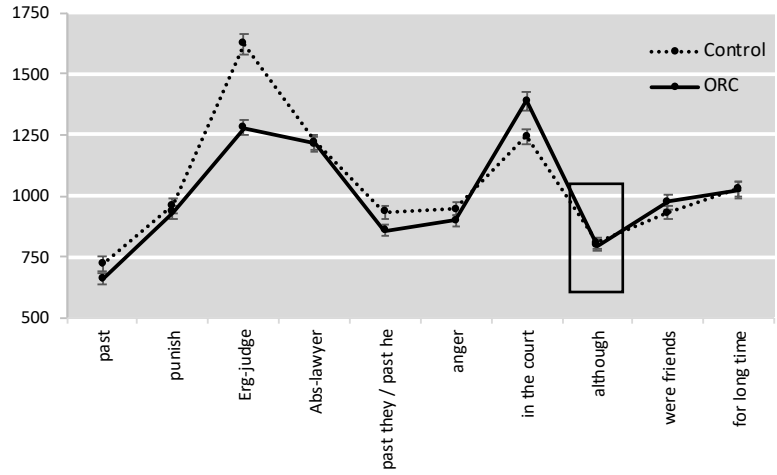

Figure 2. Region-by-region reading times for the ORC pairs.
The box indicates the critical region abs-NP. Error bars represent SEs.

4.5. Discussion

To summarize the results of the experiment, we observed a reading time slowdown in region 8 in the SRC pairs, but no similar slowdown in the ORC pairs. The slowdown in the SRC pairs suggests that there was an ORC preference in Tongan. We interpret the data to indicate that the native speakers of Tongan assume the clitic *ne* as a subject pronoun when they see "*ne*+verb" at the beginning of the RCs. This further leads them to posit a gap for the absolutive object NP. In the experimental sentence in SRC (Resumptive Pronoun) condition (10b), however, they encountered the unexpected absolutive NP, which forces them to revise the structure, rendering the reading time slowdown.

The lack of slowdown in Region 8 (adjunct clause marker) in the target condition (10d) in the ORC pairs also suggests that the native speakers of Tongan took the clitic *ne* as a subject pronoun. The reading time data show that there is no disruption due to no appearance of the absolutive case-marked NP. One might suggest, however, that the reading time slowdown observed in Region 7 (PP) in fact indicates the readers' SRC preference. Taking the ambiguous *ne* in RC to be the resumptive pronoun, the readers expect to see an absolutive case-marked NP immediately after the verb. Under this account, a longer reading time of PP can be taken as a mismatch from their expectation. At this point, it is not clear to what extent the presence of PP immediately after the verb can be a strong cue falsifying a

particular RC structure. We are certainly interested in investigating this point in future studies.

If we are on the right track, the ORC preference suggests that there is an absolutive Case advantage in Tongan, like Avar and Basque (Polinsky et al. 2012, Carreiras et al. 2010). The dependency with an absolutive case-marked NP, which bears an unmarked Case, is less costly to process than the dependency with a marked Case. This is one of the expected outcomes because the unmarkedness of the absolutive Case was observed in morphological ergative languages like Avar and Basque; the current data clearly show a similar effect in Tongan, a syntactic ergative language. The ORC preference also suggests that the structural distance in RCs (Hawkins 2004, O'Grady 1997) or the linear distance (Gibson 2000) does not play a major role in Tongan RC processing. Finally, the subject prominence was not observed unlike Chamorro or Avar. Because Avar is an ergative-absolutive language, it is not the case that ergativity by itself is not compatible with the subject advantage. We suggest that the syntactic ergativity is one key factor that suppresses the subject advantage in Tongan.

In the above discussion, we argue that an RC structure with an absolutive-gap is preferred in Tongan. However, an alternative account might be that there was a strong preference in Tongan to interpret the clause initial clitic *ne* as a subject pronoun, rather than a resumptive pronoun. Under this account, the ORC preference is not caused by a Case-related reason, but is a consequence of ambiguity resolution of the clitic *ne*. Unfortunately we do not have information about the frequency of the use of *ne* at this point, so we cannot provide a relevant discussion from that point of view. One might argue that taking the clitic *ne* as a resumptive pronoun is quite costly; but it is not so clear whether resumptive pronouns in general are costly to process in Tongan, because resumptive pronouns in Tongan are widely available, and they are more like various kinds of agreement markers, rather than "intrusive" resumptive pronouns found in English, for instance. In some future experiments, we would like to test RC structures with those factors controlled.

5. Conclusion

In this paper, we report our results obtained from a self-paced reading experiment in Tongan, a V-initial language with syntactic ergativity. Comparing the reading time data from SRC and ORC, the data suggest that native speakers of Tongan prefer the ORC structure, positing a gap in the transitive object position. This suggests that Tongan seems not to exhibit the subject advantage observed in other ergative languages like Avar, or other Austronesian languages like Chamorro. Instead, the data support that the absolutive Case advantage based on the unmarkedness of the Case in ergative languages plays a major role in processing RCs. The lack of the subject advantage may be linked to the syntactic ergativity of this language, but further investigation about the syntactic ergativity, especially about the resumptive pronouns, is needed in future.

References

Aldridge, Edith. 2004. Ergativity and word order in Austronesian languages. Doctoral dissertation, Cornell University.

Anand, Pranav, Sandra Chung, and Matthew Wagers. 2011. Widening the net: Challenges for gathering linguistic data in the digital age. NSF SBE 2020: Future research in the social, behavioral and economic sciences, URL https://www.nsf.gov/sbe/sbe_2020/.

Baayen, Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, Harald, Debra J. Davidson, and Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4): 390–412.

Baayen, Harald, and Petar Milin. 2010. Analyzing Reaction Times. *International Journal of Psychological Research* 3(2): 12–28.

Bader, Markus, and Michael Meng. 1999. Subject-object ambiguities in German embedded clauses: An across-the-board comparison. *Journal of Psycholinguistic Research* 28: 121–143.

Bates, Douglas, Martin Mächler, Benjamin M. Bolker, and Steve C. Walker. 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7, URL http://CRAN.R-project.org/package=lme4.

Carreiras, Manuel, Jon Andoni Duñabeitia, Marta Vergara, Irene de la Cruz-Pavía, and Itziar Laka. 2010. Subject relative clauses are not universally easier to process: Evidence from Basque. *Cognition* 115: 79–92.

Churchward, Maxwell. 1953. *Tongan grammar*. Oxford: Oxford University Press.

Clemens, Lauren, Jessica Coon, Pedro Mateo Pedro, Adam Milton Morgan, Maria Polinsky, Gabrielle Tandet, and Matthew Wagers. 2015. Ergativity and the complexity of extraction: A view from Mayan. *Natural Language and Linguistic Theory* 33(2): 417–467.

Comrie, Bernard. 2013. Alignment of case marking of full noun phrases. *The world atlas of language structures online*, eds. Matthew Dryer and Martin Haspelmath, Leipzig: Max Planck Institute for Evolutionary Anthropology. URL http://wals.info/.

Comrie, Bernard, and Edward L. Keenan. 1979. Noun phrase accessibility revisited. *Language* 55:649–64.

Crain, Stephen, and Janet Dean Fodor. 1985. How can grammars help parsers? Natural language parsing: Psychological, computational and theoretical perspectives, eds. David R. Dowty, Lauri Kartunnen, and Arnold M. Zwicky, 94–128. Cambridge: Cambridge University Press.

Dryer, Matthew. 2013. Order of subject, object and verb. *The world atlas of language structures online*, eds. Matthew Dryer and Martin Haspelmath, Leipzig: Max Planck Institute for Evolutionary Anthropology. URL http://wals.info/.

Frazier, Lyn, and Keith Rayner. 1982. Making and correcting errors during sentence comprehension: Eye-movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology* 14, 178–210.

García Matzar, Lolmay Pedro Oscar, and José Obispo Rodríguez Guaján. 1997. *Rukemik ri Kaqchikel chi': Gramática Kaqchikel*. Guatemala City: Cholsamaj.

Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. *Image, language, brain: Papers from the first min articulation project symposium*, eds. Alec Marantz, Yasushi Miyashita and Wayne O'Neil, 94–126. Cambridge, MA: MIT Press.

Hawkins, John. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Just, Marcel, Patricia Carpenter, and Jacqueline Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 111: 228–38.

Keenan, Edward L., and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63–99.

Keenan, Edward L., and Bernard Comrie. 1979. Data on the noun phrase accessibility hierarchy. *Language* 55(2): 333–351.

King, Jonathan, and Marcel Adam Just. 1991. Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language* 30: 580–602.

Koizumi, Masatoshi, and Kim, Jungho. 2016. Greater left inferior activation for SVO than VOS during sentence comprehension in Kaqchikel. *Frontiers in Psychology* 7: 1541, doi: 10.3389/fpsyg.2016.01541.

Koizumi, Masatoshi, Yoshiho Yasugi, Katsuo Tamaoka, Sachiko Kiyama, Jungho Kim, Juan Estevan Ajsivinac Sian, and Lolmay Pedro Oscar García Mátzar. 2014. On the (non)universality of the preference for subject-object word order in sentence comprehension: A sentence-processing study in Kaqchikel Maya. *Language* 90: 722–736.

Kwon, Nayoung, Younhyoung Lee, Peter C. Gordon, Robert Kluender, and Maria Polinsky. 2010. Cognitive and linguistic factors affecting subject/object asymmetry: An eye-tracking study of prenominal relative clauses in Korean. *Language* 86(3): 546–582.

Kwon, Nayoung, Robert Kluender, Martha Kutas, and Maria Polinsky. 2013. Subject/object processing asymmetries in Korean relative clauses: Evidence from ERP data. *Language* 89(3): 537–585.

Lau, Ellen, Michaela Socolof, Nancy Clarke, Rusudan Asatiani, and Masha Polinsky. 2018. A subject relative clause preference in a split-ergative language: ERP evidence from Georgian. Ms. University of Maryland.

Longenbaugh, Nicholas, and Maria Polinsky. 2016. The processing of long-distance dependencies in Niuean. The Proceedings of AFLA 22, ed. Henrison Hsieh, 98–120. Canberra: Asia-Pacific Linguistics.

Marantz, Alec. 1991. Case and licensing. *Proceedings of the Eastern States Conference on Linguistics (ESCOL)* 9, 234–253. The Department of Linguistics, the Ohio State University.

Norcliffe, Elisabeth, Agnieszka E. Konopka, Penelope Brown, and Stephen C. Levinson. 2015. Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience*, 30(9): 1187–1208.

O'Grady, William. 1997. *Syntactic development*. Chicago, IL: University of Chicago Press.

Otsuka, Yuko. 2000. Ergativity in Tongan. Doctoral dissertation, University of Oxford.

Polinsky, Maria, Carlos Goméz Gallo, Peter Graff, and Ekaterina Kravtchenko. 2012. Subject preference and ergativity. *Lingua* 122: 267–277.

R Development Core Team. 2014. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, URL http://www.R-project.org/.

Sauppe, Sebastian. 2016. Verbal semantics drives early anticipatory eye movements during the comprehension of verb-initial sentences. *Frontiers in Psychology* 7:95. DOI: 10.3389/fpsyg.2016.00095

Sauppe, Sebastian, Elizabeth J. Norcliffe, Agnieszka E. Konopka, Robert D. Van Valin, Jr., and Stephen C. Levinson. 2013. Dependencies first: Eye tracking evidence from sentence production in Tagalog. Proceedings of the 35th Annual Conference of the Cognitive Science Society, eds. M. Knauff, M. Pauen, N. Sebanz, and I. Wachsmuth, 1265–1270. Austin, TX: Cognitive Science Society.

Tanaka, Nozomi. 2016. An asymmetry in the acquisition of Tagalog relative clauses. Doctoral dissertation, University of Hawai'i at Mānoa.

Tollan, Rebecca. 2019. Cross-linguistic effects of subjecthood, case, and transitivity in syntax and sentence processing. Doctoral dissertation, University of Tronto.

Tollan, Rebecca, Diane Massam, and Daphna Heller. 2019. Effects of case and transitivity on processing dependencies: Evidence from Niuean. *Cognitive Science*, e12736. DOI: 10.1111/cogs.12736.

Traxler, Matthew J., Robin K. Morris, and Rachel E. Seely. 2002. Processing subjectand object relative clauses: Evidence from eye movements. *Journal of Memory and Language* 47: 69–90.

Wagers, Matt, Manual F. Borja, and Sandra Chung. 2015. The real-time comprehension of WH-dependences in a WH-agreement language. *Language* 91(1): 109–144.

Wagers, Matthew, Manual F. Borja, and Sandra Chung. 2018. Grammatical licensing and relative clause parsing in a flexible word-order language. *Cognition* 178: 207–221.

Yano, Masataka, Daichi Yasunaga, and Masatoshi Koizumi. 2017. Event-related brain indices of gap-filling processing in Kaqchikel. *Event-Related*

Potential(ERP): Methods, Outcomes, Research Insights, ed. Samuel. R. Harris, 89–122. NOVA Science Publishers.

Yasunaga, Daichi, Masataka Yano, Yoshiho Yasugi, and Masatoshi Koizumi. 2015. Is the subject-before-object preference universal? An event-related potential study in the Kaqchikel Mayan language. *Language, Cognition and Neuroscience* 30: 1209–1229.